

An Algorithm for Deciding the Summability of Bivariate Rational Functions

Qing-Hu Hou¹ and Rong-Hua Wang²

¹Center for Applied Mathematics
Tianjin University, Tianjin 300072, P.R. China
hou@nankai.edu.cn

²Center for Combinatorics, LPMC-TJKLC
Nankai University, Tianjin 300071, P.R. China
wangwang@mail.nankai.edu.cn

Abstract. Let $\Delta_x f(x, y) = f(x + 1, y) - f(x, y)$ and $\Delta_y f(x, y) = f(x, y + 1) - f(x, y)$ be the difference operators with respect to x and y . A rational function $f(x, y)$ is called summable if there exist rational functions $g(x, y)$ and $h(x, y)$ such that $f(x, y) = \Delta_x g(x, y) + \Delta_y h(x, y)$. Recently, Chen and Singer presented a method for deciding whether a rational function is summable. To implement their method in the sense of algorithms, we need to solve two problems. The first is to determine the shift equivalence of two bivariate polynomials. We solve this problem by presenting an algorithm for computing the dispersion sets of any two bivariate polynomials. The second is to solve a univariate difference equation in an algebraically closed field. By considering the irreducible factorization of the denominator of $f(x, y)$ in a general field, we present a new criterion which requires only finding a rational solution of a bivariate difference equation. This goal can be achieved by deriving a universal denominator of the rational solutions and a degree bound on the numerator. Combining these two algorithms, we can decide the summability of a bivariate rational function.

Keywords: summability, bivariate rational function, Gosper's algorithm, dispersion set.

1 Introduction

In 1978, Gosper [13] presented the celebrated algorithm which solves the problem of determining whether a given hypergeometric term is equal to the difference of another hypergeometric term. Based on Gosper's algorithm, Zeilberger [25, 26] gave a fast algorithm for proving terminating hypergeometric identities. Zeilberger's method was further extended to the multivariate case by Wilf and Zeilberger himself in [24]. Paule [20] gave an interpretation of Gosper's algorithm in terms of the greatest factorial factorizations. Chen, Paule and Saad [12] derived an easy understanding version of Gosper's algorithm by considering the convergence of the greatest common divisors of two polynomial sequences.

Other approaches to the summability of rational functions were given by Abramov [2–4], Pirastu and Strehl [22], Ash and Catoiu [8]. The key idea of these methods is to rewrite a rational function α as $\alpha = \Delta(\beta) + \gamma$, where Δ is the difference operator, β and γ are rational functions such that the denominator of γ is shift-free. Then α is summable if and only if γ is zero.

Passing from univariate to multivariate, Zeilberger's algorithm have been discussed by Zeilberger himself [7, 19], Koutschan [18], Schneider [23], Chen et. al. [11]. These algorithms are useful in practice. However, they did not provide a complete answer to the summability problem of bivariate hypergeometric terms. Only very recently, Chen and Singer [10] presented criteria for deciding the summability of *bivariate rational functions*. When applying their criteria, one will encounter two problems. The first one is how to determine whether two bivariate polynomials are shift equivalent. The second one is how to solve univariate difference equations in algebraically closed fields. The main aim of the present paper is to overcome these problems and to give an algorithm for deciding the summability of bivariate rational functions. We remark that the general question considered in this paper was raised by Andrews and Paule in [6].

For the first problem, we show that the dispersion set of two bivariate polynomials is computable. Then two polynomials are shift equivalent if and only if the dispersion set is not empty. For the second problem, we present a variation of the criteria by considering the irreducible factorization in a general field instead of an algebraically closed field. To apply the new criteria, we need only to find rational solutions of a bivariate difference equation. By a discussion similar to Gosper's algorithm, we derive a universal denominator of the rational solutions. We further derive a degree bound on the numerator of the rational solutions and thus obtain an algorithm for the new criterion. Combining these two algorithms, we finally obtain an algorithm for deciding the summability of bivariate rational functions.

The paper is organized as follows. In Section 2, we give an algorithm for computing the dispersion set of two bivariate polynomials. In Section 3, we first reduce the summability of a general rational function to that of a rational function whose denominator is a power of an irreducible polynomial. Then we present a criterion on the summability of this special kind of rational functions. This criterion reduces the summability problem to the problem of finding rational solutions of a bivariate difference equation. In Section 4, we give an algorithm for solving the difference equation.

Throughout the paper, we take \mathbb{Q} , the field of rational numbers, as the ground field. It should be mentioned that the discussions work also for other fields, such as the extension field $\mathbb{Q}(\alpha_1, \dots, \alpha_r)$ where $\alpha_1, \dots, \alpha_r$ are either algebraic or transcendental over \mathbb{Q} .

We follow the notations used in [10]. Let $f(x, y) \in \mathbb{Q}(x, y)$ be a bivariate rational function. The shift operators σ_x and σ_y are given by

$$\sigma_x f(x, y) = f(x + 1, y) \quad \text{and} \quad \sigma_y f(x, y) = f(x, y + 1).$$

A function $f \in \mathbb{Q}(x, y)$ is said to be (σ_x, σ_y) -summable if there exist two rational functions $g, h \in \mathbb{Q}(x, y)$ such that

$$f = \sigma_x g - g + \sigma_y h - h.$$

2 Dispersion set and shift equivalence

Let \mathbb{Z} denote the set of integers. Recall that given two univariate polynomials, say $f(x)$ and $g(x)$, their *dispersion set* is defined by

$$\text{Disp}_x(f, g) = \{n \in \mathbb{Z} \mid f(x) = g(x + n)\}.$$

It is known that unless f and g are the same constant polynomial, the dispersion set $\text{Disp}_x(f, g)$ is finite and is computable. For the algorithm, see [21, page 79]. We can extend this concept to the bivariate case.

Definition 2.1. Let f, g be two bivariate polynomials in $\mathbb{Q}[x, y]$ and σ_x, σ_y be the shift operators. The dispersion set of f and g is defined by

$$\text{Disp}(f, g) = \{(m, n) \in \mathbb{Z}^2 \mid f = \sigma_x^m \sigma_y^n g\}.$$

If $\text{Disp}(f, g)$ is not empty, we say f and g are shift equivalent.

In particular, when $f = \sigma_x^m g$ (resp. $f = \sigma_y^n g$), we say f, g in the same σ_x -orbit (resp. σ_y -orbit), denoted by $f \sim_x g$ and $f \sim_y g$ respectively.

We remark that testing shift equivalence over fields have been considered by Grigoriev and Karpinski [14–16]. More precisely, they gave algorithms to find shifts $(\alpha_1, \dots, \alpha_r) \in F^r$ such that

$$f(x_1 + \alpha_1, \dots, x_r + \alpha_r) = g(x_1, \dots, x_r),$$

where F is a field and $f, g \in F[x_1, \dots, x_r]$. Instead of considering shifts over a field, we focus on integer shifts, i.e., $m, n \in \mathbb{Z}$.

In the univariate case, the dispersion set of any two polynomials is computable. The following theorem shows that the dispersion set is also computable in the bivariate case.

Theorem 2.2. Let $f, g \in \mathbb{Q}[x, y]$ be two polynomials. Then we can determine the dispersion set $\text{Disp}(f, g)$.

Proof. Since the shift operators σ_x and σ_y preserve the degree, we get that $\text{Disp}(f, g) = \emptyset$ unless $\deg_x f = \deg_x g$.

When $f = 0$ or $\deg_x(f) = 0$, the computation of $\text{Disp}(f, g)$ reduces to the univariate case. More precisely, we have

$$\text{Disp}(f, g) = \mathbb{Z} \times \text{Disp}_y(f, g).$$

Now assume that $\deg_x f = d > 0$ and write f, g as

$$f = \sum_{k=0}^d a_k(y)x^k, \quad g = \sum_{k=0}^d b_k(y)x^k.$$

Suppose that $(m, n) \in \text{Disp}(f, g)$. By comparing the leading coefficient with respect to x , we see that n falls in the dispersion set

$$\mathcal{N} = \text{Disp}_y(a_d(y), b_d(y)).$$

If \mathcal{N} is a finite set, we then have

$$\text{Disp}(f, g) = \bigcup_{n_0 \in \mathcal{N}} \text{Disp}_x(f(x, y), g(x, y + n_0)) \times \{n_0\}.$$

Otherwise, we may assume $a_d(y) = b_d(y) = c$, where c is a non-zero constant. By comparing the second leading coefficient with respect to x , we see that

$$a_{d-1}(y) = d \cdot c \cdot m + b_{d-1}(y + n). \quad (2.1)$$

According to the degree of $a_{d-1}(y)$ in variable y , there are three cases.

Case 1. $\deg a_{d-1}(y) > 1$. Then $\text{Disp}(f, g) = \emptyset$ unless the leading term of $a_{d-1}(y)$ and that of $b_{d-1}(y)$ coincide. Assume

$$a_{d-1}(y) = \sum_{j=0}^h p_j y^j \quad \text{and} \quad b_{d-1}(y) = \sum_{j=0}^h q_j y^j.$$

By comparing the coefficients of y^{h-1} in the expansions of $a_{d-1}(y)$ and $b_{d-1}(y + n)$, we see that n is uniquely determined by

$$h q_h n + q_{h-1} = p_{h-1}. \quad (2.2)$$

Suppose that n_0 is an integer solution of (2.2). We then have

$$\text{Disp}(f, g) = \text{Disp}_x(f(x, y), g(x, y + n_0)) \times \{n_0\}.$$

Case 2. $\deg a_{d-1}(y) = 1$. We also have $\text{Disp}(f, g) = \emptyset$ unless the leading term of $a_{d-1}(y)$ and that of $b_{d-1}(y)$ coincide. Assume

$$a_{d-1}(y) = p_1 y + p_0 \quad \text{and} \quad b_{d-1}(y) = p_1 y + q_0.$$

Then (2.1) leads to

$$(d \cdot c) \cdot m + p_1 \cdot n = p_0 - q_0, \quad (2.3)$$

which is a linear Diophantine equation in unknowns m, n . Either there is no solution, or the solutions are of the form

$$m = ut + v, \quad \text{and} \quad n = u't + v',$$

where u, v, u', v' are explicit integers and t runs over \mathbb{Z} . Now by setting all coefficients of x, y in the expansion of $f(x, y) - g(x + ut + v, y + u't + v')$ to be zeros, we arrive at a system of polynomial equations in t . The set of integer solutions of each equation is computable (see, for example [21, page 79]). The final dispersion set of f and g is the intersection of these solution sets.

Case 3. $\deg a_{d-1}(y) = 0$ or $a_{d-1}(y) = 0$. If $\deg_y b_{d-1}(y) > 0$, we then have $\text{Disp}(f, g) = \emptyset$. Otherwise, m is uniquely determined by (2.1). Suppose m_0 is an integer solution of (2.1), we have

$$\text{Disp}(f, g) = \{m_0\} \times \text{Disp}_y(f(x, y), g(x + m_0, y)).$$

This completes the proof. ■

Based on the proof as above, we can describe an algorithm for computing the dispersion set of two polynomials in $\mathbb{Q}[x, y]$.

Algorithm DispSet

Input: Two polynomials $f = \sum_{k=0}^{d_1} a_k(y)x^k$ and $g = \sum_{k=0}^{d_2} b_k(y)x^k$.

Output: The dispersion set $\text{Disp}(f, g)$.

1. If $d_1 \neq d_2$, return \emptyset . Else set $d = d_1 = d_2$.
2. If $d \leq 0$, return the set $\mathbb{Z} \times \text{Disp}_y(f, g)$. Else continue the following steps.
3. If $\deg a_d(y) > 0$, compute $\mathcal{N} = \{n \mid a_d(y) = b_d(y + n)\}$ and for each $n_0 \in \mathcal{N}$, compute the set S_{n_0} of integers m such that $f = \sigma_x^m \sigma_y^{n_0} g$. Return the set

$$\bigcup_{n_0 \in \mathcal{N}} S_{n_0} \times \{n_0\}.$$

Else set $a_d(y) = c$ and continue the following steps.

4. If $\deg_y a_{d-1}(y) > 1$, compute the unique n_0 according to (2.2). If n_0 is an integer, then return $\text{Disp}_x(f(x, y), g(x, y + n_0)) \times \{n_0\}$. Else return \emptyset .
5. If $\deg_y a_{d-1}(y) = 1$. If the leading terms of $a_{d-1}(y)$ and $b_{d-1}(y)$ are different, then return \emptyset . Else solve the linear Diophantine equation (2.3). Suppose that the solutions are of the form

$$m = ut + v \quad \text{and} \quad n = u't + v'.$$

Substituting m by $ut + v$ and n by $u't + v'$ in $f = \sigma_x^m \sigma_y^n g$ and comparing the coefficients of each power of x and y to get a system of polynomial equations in t . Return all integer solutions if there are. Else return \emptyset .

6. If $\deg_y a_{d-1}(y) = 0$ or $a_{d-1}(y) = 0$. If $\deg_y b_{d-1}(y) > 0$ then return \emptyset . Else compute the unique m_0 satisfying (2.1). If m_0 is not an integer, then return \emptyset . Else return the set

$$\{m_0\} \times \text{Disp}_y(f(x, y), g(x + m_0, y)).$$

The following is an example which shows how to determine the shift equivalence of any two given bivariate polynomials.

Example 2.3. *Let*

$$f = 2x^2 + 2xy + y^2 + y + 1 \quad \text{and} \quad g = 2x^2 + 2xy + y^2 + 2x + y + 1.$$

We try to determine whether f and g are shift equivalent according to the proof of Theorem 2.2. Rewrite f, g as

$$f = 2x^2 + (2y)x + (y^2 + y + 1), \quad \text{and} \quad g = 2x^2 + (2y + 2)x + (y^2 + y + 1).$$

It's easy to check that this meets Case 2 in the proof. Thus m, n satisfy the linear equation $2m + n = -1$ whose solutions are

$$m = t \quad \text{and} \quad n = -2t - 1, \quad t \in \mathbb{Z}.$$

Now by setting all coefficients of x, y in the expansion of $f(x, y) - g(x + t, y - 2t - 1)$ to be zeros, we obtain an integer solution $t = -1$. It means that $f(x, y) = g(x - 1, y + 1)$ and thus f, g are shift equivalent.

3 Summability criterion

As stated in the introduction, one can decompose a univariate rational function α into the form $\alpha = \Delta\beta + \gamma$. The goal of this section is to introduce a bivariate variant of such additive decomposition and thus reduce the bivariate summability problem of a general rational function to that of a rational function whose denominator is a power of an irreducible polynomial. We then present a criterion for the summability of this kind of special rational functions.

Let $f \in \mathbb{Q}(x, y)$ be a bivariate rational function. Assume that the irreducible factorization of the denominator $D(x, y)$ of $f(x, y)$ is

$$D(x, y) = \prod_{i=1}^m d_i^{n_i}(x, y),$$

where $d_i(x, y)$ are irreducible polynomials and n_i are positive integers. Viewing f as a rational function of y over the field $\mathbb{Q}(x)$, we have the partial fraction decomposition

$$f = P + \sum_{i=1}^m \sum_{j=1}^{n_i} \frac{a_{i,j}}{d_i^j}, \quad (3.1)$$

where $P \in \mathbb{Q}(x)[y]$, $a_{i,j} \in \mathbb{Q}(x)[y]$ and $\deg_y(a_{i,j}) < \deg_y(d_i)$. It is well known that the polynomial P is the difference of a polynomial.

Now suppose that $d_i(x, y) = d_k(x + m, y + n)$ for some index $i \neq k$. Then we have

$$\frac{a_{i,j}}{d_i^j} = \sigma_x(g) - g + \sigma_y(h) - h + \frac{\sigma_x^{-m} \sigma_y^{-n}(a_{i,j})}{d_k^j},$$

where

$$g = \begin{cases} \sum_{\ell=0}^{m-1} \frac{\sigma_x^{\ell-m}(a_{i,j})}{\sigma_x^{\ell} \sigma_y^n(d_k^j)}, & \text{if } m \geq 0, \\ - \sum_{\ell=0}^{-m-1} \frac{\sigma_x^{\ell}(a_{i,j})}{\sigma_x^{m+\ell} \sigma_y^n(d_k^j)}, & \text{if } m < 0, \end{cases}$$

and

$$h = \begin{cases} \sum_{\ell=0}^{n-1} \frac{\sigma_y^{\ell-n} \sigma_x^{-m}(a_{i,j})}{\sigma_y^{\ell}(d_k^j)}, & \text{if } n \geq 0, \\ - \sum_{\ell=0}^{-n-1} \frac{\sigma_y^{\ell} \sigma_x^{-m}(a_{i,j})}{\sigma_y^{n+\ell}(d_k^j)}, & \text{if } n < 0. \end{cases}$$

Repeating the above transformation, we arrive at the following decomposition.

Lemma 3.1. *For a rational function $f \in \mathbb{Q}(x, y)$, we can decompose it into the form*

$$f = \Delta_x(g) + \Delta_y(h) + r,$$

where $g, h \in \mathbb{Q}(x, y)$ and r is of the form

$$r = \sum_{i=1}^m \sum_{j=1}^{n_i} \frac{a_{i,j}(x, y)}{d_i^j(x, y)}, \quad (3.2)$$

with $a_{i,j} \in \mathbb{Q}(x)[y]$, $\deg_y(a_{i,j}) < \deg_y(d_i)$, $d_i \in \mathbb{Q}[x,y]$ are irreducible polynomials, and d_i and $d_{i'}$ are not shift equivalent for any $1 \leq i \neq i' \leq m$.

From Lemma 3.1, we see that f is (σ_x, σ_y) -summable if and only if r is (σ_x, σ_y) -summable. The following lemma shows that the summability of r is equivalent to the summability of each summand of r .

Lemma 3.2. *Let $r \in \mathbb{Q}(x,y)$ be of the form (3.2). Then r is (σ_x, σ_y) -summable if and only if $\frac{a_{i,j}(x,y)}{d_i^j(x,y)}$ is (σ_x, σ_y) -summable for all $1 \leq i \leq m$ and $1 \leq j \leq n_i$.*

Proof. The sufficiency follows from the linearity of the difference operators Δ_x and Δ_y . It suffices to prove the necessity. Assume that r is (σ_x, σ_y) -summable, then there exist $g, h \in \mathbb{Q}(x,y)$ such that $r = \sigma_x(g) - g + \sigma_y(h) - h$. We can always decompose g, h as

$$g = \frac{A_1}{D_1} + \frac{A_2}{D_2} \quad \text{and} \quad h = \frac{B_1}{C_1} + \frac{B_2}{C_2},$$

where $A_i, B_i, C_i, D_i (i = 1, 2)$ are polynomials in y over $\mathbb{Q}(x)$, $\deg_y(A_1) < \deg_y(D_1)$, $\deg_y(B_1) < \deg_y(C_1)$, D_1 (resp. C_1) contains only irreducible factors that are shift equivalent to d_i , while D_2 (resp. C_2) contains no such factors. Let $r_i = \sum_{j=1}^{n_i} \frac{a_{i,j}(x,y)}{d_i^j(x,y)}$. We then have

$$r_i - \left(\sigma_x \frac{A_1}{D_1} - \frac{A_1}{D_1} + \sigma_y \frac{B_1}{C_1} - \frac{B_1}{C_1} \right) = \sigma_x \frac{A_2}{D_2} - \frac{A_2}{D_2} + \sigma_y \frac{B_2}{C_2} - \frac{B_2}{C_2} - \sum_{j \neq i} r_j.$$

Note that σ_x, σ_y preserve the (σ_x, σ_y) -equivalence. Therefore, we have

$$r_i = \sigma_x \frac{A_1}{D_1} - \frac{A_1}{D_1} + \sigma_y \frac{B_1}{C_1} - \frac{B_1}{C_1},$$

which means r_i is (σ_x, σ_y) -summable.

By the same observation as in [10, Page 330], we see that σ_x and σ_y preserve the multiplicities of the fractions $a_{i,j}/d_i^j$. This implies that r_i is (σ_x, σ_y) -summable if and only if each summand $a_{i,j}/d_i^j$ is (σ_x, σ_y) -summable. This concludes the proof. \blacksquare

Now we only need to study the summability problem of rational functions of the form a/d^j , where $d \in \mathbb{Q}[x,y]$ is irreducible, $a \in \mathbb{Q}(x)[y]$, and $\deg_y(a) < \deg_y(d)$. For this kind of rational functions, we have the following criterion for their summability.

Theorem 3.3. *Let $f = \frac{a(x,y)}{d^j(x,y)}$, where $d(x,y) \in \mathbb{Q}[x,y]$ is an irreducible polynomial, $a \in \mathbb{Q}(x)[y]$ is non-zero and $\deg_y(a) < \deg_y(d)$. Then f is (σ_x, σ_y) -summable if and only if*

(1) *there exist integers t, ℓ with $t \neq 0$ such that*

$$\sigma_x^t d(x,y) = \sigma_y^\ell d(x,y), \tag{3.3}$$

(2) *for the smallest positive integer t such that (3.3) holds, we have*

$$a = \sigma_x^t \sigma_y^{-\ell} p - p, \tag{3.4}$$

for some $p \in \mathbb{Q}(x)[y]$ with $\deg_y(p) < \deg_y(d)$.

We can adapt the argument used in [10, Theorem 3.7] to complete the proof of Theorem 3.3. The details are elaborated in the appendix.

The criterion (3.3) can be tested by computing the dispersion set $\text{Disp}(d, d)$. In the next section, we will give an algorithm for solving the equation (3.4). Then combining Lemma 3.1, Lemma 3.2 and Theorem 3.3, we will obtain an algorithm for determining whether a bivariate rational function is summable.

4 Rational solutions of the difference equations

Let d_0 be a positive integer and u be a polynomial in y over $\mathbb{Q}(x)$ with $\deg_y(u) < d_0$. In this section, we present a method of finding solutions $p \in \mathbb{Q}(x)[y]$ with $\deg_y(p) < d_0$ to the following difference equation

$$u = \sigma_x^m \sigma_y^{-n} p - p, \quad (4.1)$$

where m, n are given integers and $m > 0$.

Noting that $\deg_y(p) < d_0$, we may assume

$$p = p_0(x) + p_1(x)y + \cdots + p_{d_0-1}(x)y^{d_0-1}.$$

Then comparing the coefficients of each power of y on both sides of (4.1), we obtain a system of linear difference equations in $p_i(x)$. Abramov-Barkatou [5] and Abramov-Khmelnov [1] presented algorithms for solving such systems.

We will rewrite p as the ratio $c(x, y)/d(x)$ and estimate the denominator $d(x)$ directly. Then we give an upper bound on the x -degree of the numerator $c(x, y)$ and thus solve for p by the method of undetermined coefficients.

Assume that $u = a(x, y)/b(x)$, where a, b are polynomials in x and y . We notice that one can give an estimation of $d(x)$ by using the convergence argument introduced by Chen, Paule and Saad [12]. More precisely, we have

$$d(x) \mid \gcd(b(x)b(x+m)\cdots, b(x-m)b(x-2m)\cdots).$$

Note also that one can give an estimation by using an argument similar to [17]. Here we give another estimation based on Gosper representation [21, page 80].

Rewrite Equation (4.1) as

$$a(x, y) = \frac{b(x)}{b(x+m)} \sigma_x^m \sigma_y^{-n} (b(x)p(x, y)) - b(x)p(x, y). \quad (4.2)$$

Let

$$\frac{b(x)}{b(x+m)} = \frac{A(x)}{B(x)} \frac{C(x+m)}{C(x)}, \quad (4.3)$$

be the Gosper representation. That is,

$$\gcd(A(x), B(x+hm)) = 1, \quad \forall h \in \mathbb{N}. \quad (4.4)$$

Then the denominator of p can be given by the following theorem.

Theorem 4.1. *Let $(A(x), B(x), C(x))$ be the Gosper representation of $\frac{b(x)}{b(x+m)}$, and $b(x)p(x, y)$ be a rational solution of (4.2). Then $b(x)p(x, y)$ must be of the form*

$$b(x)p(x, y) = \frac{B(x-m)p_1(x, y)}{C(x)},$$

where $p_1(x, y)$ is a polynomial in both x and y .

Proof. Assume that

$$b(x)p(x, y) = \frac{g(x, y)}{q(x)C(x)},$$

where $g(x, y) \in \mathbb{Q}[x, y]$, $q(x) \in \mathbb{Q}[x]$ is a monic polynomial and $(q(x), g(x, y)) = 1$. According to Equation (4.2), we deduce that

$$a(x, y)B(x)C(x)q(x)q(x+m) = A(x)g(x+m, y-n)q(x) - B(x)g(x, y)q(x+m). \quad (4.5)$$

It's easy to check that

$$q(x) \mid g(x, y)B(x)q(x+m).$$

Since $(q(x), g(x, y)) = 1$, we obtain

$$q(x) \mid B(x)q(x+m).$$

Using this divisibility repeatedly, we can get

$$q(x) \mid B(x)B(x+m) \cdots B(x+(r-1)m)q(x+rm).$$

When $r > \max \text{Disp}_x(q(x), q(x))$, we have $(q(x), q(x+rm)) = 1$, and thus

$$q(x) \mid B(x)B(x+m) \cdots B(x+(r-1)m).$$

From Equation (4.5), we also derive that

$$q(x+m) \mid g(x+m, y-n)A(x)q(x).$$

By a similar discussion, we arrive at

$$q(x) \mid A(x-m)A(x-2m) \cdots A(x-rm).$$

By the definition of Gosper representation, we know that $\gcd(A(x), B(x+hm)) = 1$ for any $h \in \mathbb{N}$. Thus the only opportunity for $q(x)$ is $q(x) = 1$.

When $q(x) = 1$, Equation (4.5) will be reduced to

$$a(x, y)B(x)C(x) = A(x)g(x+m, y-n) - B(x)g(x, y).$$

It's easy to see that

$$B(x) \mid A(x)g(x+m, y-n),$$

and hence $B(x) \mid g(x+m, y-n)$. Setting $g(x, y) = B(x-m)p_1(x, y)$ concludes the proof. ■

Substituting (4.3) and $b(x)p(x, y) = \frac{B(x-m)p_1(x, y)}{C(x)}$ into (4.2), we obtain

$$a(x, y)C(x) = A(x)p_1(x + m, y - n) - B(x - m)p_1(x, y). \quad (4.6)$$

Notice that $\deg_y(p_1) < d_0$. Therefore, in order to solve for $p_1(x, y)$, it suffices to find an upper bound on $\deg_x(p_1)$.

From the Gosper representation (4.3), we see that $\deg_x(A) = \deg_x(B)$ and their leading coefficients coincide. Now we write

$$\begin{aligned} A(x) &= \sum_{k=0}^{d_1} a_k x^k, & B(x - m) &= \sum_{k=0}^{d_1} b_k x^k, & C(x) &= \sum_{k=0}^{d_2} c_k x^k, \\ a(x, y) &= \sum_{k=0}^{d_3} \alpha_k(y) x^k, & p_1(x, y) &= \sum_{k=0}^{d_4} p_k(y) x^k. \end{aligned}$$

Theorem 4.2. *Suppose $A(x)$, $B(x)$ and $C(x)$ are given in (4.3), $\deg_y(a) < d_0$ and they have the above expansions. If $p_1(x, y)$ is a polynomial that satisfies (4.6), then*

$$d_4 \leq \max \left\{ d_2 + d_3 - d_1 + d_0, \frac{b_{d_1-1} - a_{d_1-1}}{ma_{d_1}} + d_0 - 1 \right\}.$$

Proof. There are two cases concerning n :

Case 1. $n = 0$.

Since the degrees and the leading coefficients of $A(x)$ and $B(x - m)$ coincide, the leading term of the right hand of (4.6) is canceled. By considering the second leading term, we encounter two cases.

Case 1a. The second leading term is not canceled. We then have

$$d_4 = d_2 + d_3 - d_1 + 1.$$

Case 1b. The second leading term is also canceled. We must have

$$a_{d_1}(p_{d_4}(y)d_4m + p_{d_4-1}(y)) + a_{d_1-1}p_{d_4}(y) = a_{d_1}p_{d_4-1}(y) + b_{d_1-1}p_{d_4}(y),$$

which leads to

$$d_4 = \frac{b_{d_1-1} - a_{d_1-1}}{ma_{d_1}}.$$

Case 2. $n \neq 0$.

Starting from $i = 0$, we consider whether the $(i + 1)$ th leading term of the right hand of (4.6) is canceled consequently.

For $i = 0$, we have the cases 2a and 2b.

Case 2a. The leading term is not canceled. We have

$$d_4 = d_2 + d_3 - d_1.$$

Case 2b. The leading term is canceled. Then we have

$$p_{d_4}(y - n) = p_{d_4}(y),$$

which implies that $p_{d_4}(y)$ is a constant.

In general, we have the cases $2a_i$ and $2b_i$.

Case 2a_i. The $(i + 1)$ th leading term is not canceled. Then we have

$$d_4 = d_2 + d_3 - d_1 + i.$$

Case 2b_i. The $(j + 1)$ th leading term are all canceled for $j = 0, 1, 2, \dots, i$. We claim that $\deg_y(p_{d_4-j}) \leq j$ for $j = 0, 1, 2, \dots, i$. When $i = 0$, the claim holds by the discussion in *Case 2b*. Suppose we have known that $\deg_y(p_{d_4-j}) \leq j$ for $j = 0, 1, \dots, i - 1$. Now we consider the induction step from $i - 1$ to i . By the condition that the $(i + 1)$ th leading term is canceled, we have

$$\begin{aligned} & a_{d_1} (p_{d_4-i}(y - n) - p_{d_4-i}(y)) \\ &= b_{d_1-1} p_{d_4-i+1}(y) + \dots + b_{d_1-i} p_{d_4} \\ & \quad - a_{d_1} \left(p_{d_4} \binom{d_4}{i} m^i + p_{d_4-1}(y - n) \binom{d_4-1}{i-1} m^{i-1} + \dots + p_{d_4-i+1}(y - n) (d_4 - i + 1) \right) \\ & \quad - a_{d_1-1} \left(p_{d_4} \binom{d_4}{i-1} m^{i-1} + p_{d_4-1}(y - n) \binom{d_4-1}{i-2} m^{i-2} + \dots + p_{d_4-i+1}(y - n) \right) \\ & \quad - a_{d_1-2} \left(p_{d_4} \binom{d_4}{i-2} m^{i-2} + p_{d_4-1}(y - n) \binom{d_4-1}{i-3} m^{i-3} + \dots + p_{d_4-i+2}(y - n) \right) \\ & \quad - \dots - a_{d_1-i} p_{d_4}. \end{aligned} \tag{4.7}$$

Since $\deg_y(p_{d_4-j}) \leq j$ for $j = 0, \dots, i - 1$, we know that $\deg_y(p_{d_4-i}) \leq i$. This proves the claim.

Now consider the above process. Since $\deg_y(p_1(x, y)) < d_0$, there exists an integer $i \leq d_0$ such that $\deg_y(p_{d_4-i}) < i$. Without loss of generality, we assume that i is the smallest such integer. If we are in case $2a_i$, we will get an upper bound

$$d_4 = d_2 + d_3 - d_1 + i \leq d_2 + d_3 - d_1 + d_0.$$

Otherwise, we have

$$\deg_y(p_{d_4-i}(y - n) - p_{d_4-i}(y)) \leq i - 2,$$

which means that the coefficient of y^{i-1} on the right hand side of (4.7) is canceled. Thus,

$$b_{d_1-1} - a_{d_1} (d_4 - i + 1) m - a_{d_1-1} = 0,$$

leading to the upper bound:

$$d_4 = \frac{b_{d_1-1} - a_{d_1-1}}{m a_{d_1}} + i - 1 \leq \frac{b_{d_1-1} - a_{d_1-1}}{m a_{d_1}} + d_0 - 1.$$

This completes the proof. ■

Now we are ready to describe an algorithm for testing the summability of a rational function $f(x, y)$ of the form $f = \frac{a(x, y)}{b(x)d^j(x, y)}$.

Algorithm SumTest:

Input: A rational function $f = \frac{a(x, y)}{b(x)d^j(x, y)}$.

Output: True, if f is (σ_x, σ_y) -summable and return p such that Equation (4.1) holds; False, otherwise.

1. Compute $\text{Disp}(d(x, y), d(x, y))$ by algorithm *DispSet*. If it is the set $\{(0, 0)\}$, then return False. Otherwise, let (m, n) be the element in the dispersion set such that m is the minimum positive integer.

2. Compute the Gosper representation $(A(x), B(x), C(x))$ of $\frac{b(x)}{b(x+m)}$.

3. Set

$$d_0 = \deg_y(d(x, y)), \quad d_1 = \deg_x(A(x)), \quad d_2 = \deg_x(C(x)), \quad \text{and} \quad d_3 = \deg_x(a(x, y)).$$

Suppose

$$\begin{aligned} a_{d_1} &= \text{coeff}(A(x), x, d_1), \\ a_{d_1-1} &= \text{coeff}(A(x), x, d_1 - 1), \\ b_{d_1-1} &= \text{coeff}(B(x - m), x, d_1 - 1), \end{aligned}$$

where $\text{coeff}(f(x), x, m)$ denotes the coefficient of x^m in the expansion of $f(x)$.

4. Let $d_4 = \max\{d_2 + d_3 - d_1 + d_0, \left\lfloor \frac{b_{d_1-1} - a_{d_1-1}}{ma_{d_1}} \right\rfloor + d_0 - 1\}$. Set

$$p_1(x, y) = \sum_{i=0}^{d_4} \sum_{j=0}^{d_0-1} c_{i,j} x^i y^j,$$

and plug it into

$$a(x, y)C(x) = A(x)p_1(x + m, y - n) - B(x - m)p_1(x, y).$$

By comparing all the coefficients of x and y on both sides, we determine whether the above equation has a solution $p_1(x, y)$. If it has no solution, then return False. Otherwise, return a solution $p_1(x, y)$ and $p(x, y) = \frac{B(x-m)p_1(x, y)}{b(x)C(x)}$.

Finally, we give some examples to illustrate how to use our criterion for deciding the summability of some rational functions.

Example 4.3. *Let*

$$f(x, y) = -\frac{(x + y + 4)}{(x^2 + 2x + 2xy - 1 + 2y + y^2)(x^2 + 2xy + y^2 - 2)}.$$

Denote

$$d(x, y) = x^2 + 2xy + y^2 - 2 = (x + y)^2 - 2.$$

By computing the dispersion set, we find that

$$x^2 + 2x + 2xy - 1 + 2y + y^2 = d(x + 1, y).$$

By partial fraction decomposition, we derive that

$$f(x, y) = \sum_{l=0}^1 \frac{a_l(x, y)}{d(x + l, y)},$$

where

$$a_0(x, y) = -x - y, \quad a_1(x, y) = x + y + 2.$$

Using (σ_x, σ_y) -reduction, we can write $f(x, y)$ as

$$f(x, y) = \Delta_x(g_1) + r(x, y), \quad (4.8)$$

where

$$g_1(x, y) = \frac{x + y + 1}{(x + y)^2 - 2} \quad \text{and} \quad r(x, y) = \frac{1}{(x + y)^2 - 2}.$$

It is easy to see that $\sigma_x d(x, y) = \sigma_y d(x, y)$. What left now is to check whether there exists a polynomial $p(x, y)$ such that

$$1 = \sigma_x \sigma_y^{-1} p(x, y) - p(x, y). \quad (4.9)$$

The x -degree bound and y -degree bound of p are 1 and 2 respectively. By the method of undetermined coefficients, we find a solution $p(x, y) = -y - 1$. From the proof of Lemma 4.5, we find out

$$r(x, y) = \sigma_x g_2(x, y) - g_2(x, y) + \sigma_y h(x, y) - h(x, y),$$

where

$$g_2(x, y) = \frac{-y - 1}{(x + y)^2 - 2}, \quad h(x, y) = \frac{y}{(x + y)^2 - 2}.$$

Substituting into (4.8), we finally derive that

$$f(x, y) = \sigma_x g(x, y) - g(x, y) + \sigma_y h(x, y) - h(x, y),$$

where

$$g(x, y) = \frac{x}{(x + y)^2 - 2}, \quad \text{and} \quad h(x, y) = \frac{y}{(x + y)^2 - 2}.$$

Example 4.4. Let

$$f(x, y) = \frac{x^2 + x^2 y + y^2 + 1}{(x^2 + y^2)(x^3 + 2xy + xy^2 + y^3)}.$$

We can decompose it into

$$f(x, y) = \frac{x^5 + x^4 - 2x + x^2 - 2x^3 + x^4 y - xy - y^2}{(x^3 + 2xy + xy^2 + y^3)x^3(x - 2)} + \frac{-x^4 + y}{(x^2 + y^2)x^3(x - 2)}.$$

Note that

$$\sigma_x^m(x^2 + y^2) \neq \sigma_y^n(x^2 + y^2), \quad \text{for any } (m, n) \neq (0, 0).$$

Then Theorem 3.3 implies that $\frac{-x^4 + y}{(x^2 + y^2)x^3(x - 2)}$ is not (σ_x, σ_y) -summable, which leads to the result that $f(x, y)$ is not (σ_x, σ_y) -summable in $\mathbb{Q}(x, y)$.

Acknowledgments. This work was supported by the 973 Project, the PCSIRT Project of the Ministry of Education and the National Science Foundation of China.

References

- [1] S. Abramov and D. Khmelnov. Denominators of rational solutions of linear difference systems of an arbitrary order. *Programming and Computer Software*, 38(2):84–91, 2012.
- [2] S. A. Abramov. The rational component of the solution of a first order linear recurrence relation with rational right hand side. *Ž. Vyčisl. Mat. i Mat. Fiz.*, 15(4):1035–1039, 1090, 1975.
- [3] S. A. Abramov. Indefinite sums of rational functions. In *ISSAC '95: Proceedings of the 1995 International Symposium on Symbolic and Algebraic Computation*, pages 303–308, New York, NY, USA, 1995. ACM.
- [4] S. A. Abramov. Rational solutions of linear difference and q -difference equations with polynomial coefficients. In *ISSAC '95: Proceedings of the 1995 International Symposium on Symbolic and Algebraic Computation*, pages 285–289, New York, NY, USA, 1995. ACM.
- [5] S. A. Abramov and M. A. Barkatou. Rational solutions of first order linear difference systems. In *ISSAC '98: Proceedings of the 1998 International Symposium on Symbolic and Algebraic Computation*, pages 124–131, New York, NY, USA, 1998. ACM.
- [6] G. E. Andrews and P. Paule. Some questions concerning computer-generated proofs of a binomial double-sum identity. *J. Symbolic Comput.*, 16(2):147–153, 1993.
- [7] M. Apagodu and D. Zeilberger. Multi-variable Zeilberger and Almkvist-Zeilberger algorithms and the sharpening of Wilf-Zeilberger theory. *Adv. in Appl. Math.*, 37(2):139–152, 2006.
- [8] J. M. Ash and S. Catoiu. Telescoping, rational-valued series, and zeta functions. *Trans. Amer. Math. Soc.*, 357(8):3339–3358, 2005.
- [9] S. Chen and M. F. Singer. Residues and telescopers for bivariate rational functions. *Adv. in Appl. Math.*, 49(2):111–133, 2012.
- [10] S. Chen and M. F. Singer. On the summability of bivariate rational functions. *J. Algebra*, 409:320–343, 2014.
- [11] W. Y. C. Chen, Q.-H. Hou, and Y.-P. Mu. A telescoping method for double summations. *J. Comput. Appl. Math.*, 196(2):553–566, 2006.
- [12] W. Y. C. Chen, P. Paule, and H. L. Saad. Converging to Gosper’s algorithm. *Adv. in Appl. Math.*, 41(3):351–364, 2008.
- [13] R. W. Gosper. Decision procedure for indefinite hypergeometric summation. *Proc. Natl. Acad. Sci. U.S.A.*, 75(1):40–42, 1978.
- [14] D. Grigoriev. Testing shift-equivalence of polynomials using quantum machines. In *ISSAC '96: Proceedings of the 1996 International Symposium on Symbolic and Algebraic Computation*, pages 49–54, New York, NY, USA, 1996. ACM.

- [15] D. Grigoriev. Testing shift-equivalence of polynomials by deterministic, probabilistic and quantum machines. *Theoret. Comput. Sci.*, 180(1-2):217–228, 1997.
- [16] D. Grigoriev and M. Karpinski. A zero-test and an interpolation algorithm for the shifted sparse polynomials. In *Applied algebra, algebraic algorithms and error-correcting codes*, volume 673 of *Lecture Notes in Comput. Sci.*, pages 162–169. Springer, Berlin, 1993.
- [17] Q.-H. Hou and Y.-P. Mu. Minimal universal denominators for linear difference equations. *J. Difference Equ. Appl.*, 17:977–986, 2011.
- [18] C. Koutschan. A fast approach to creative telescoping. *Math. Comput. Sci.*, 4(2-3):259–266, 2010.
- [19] M. Mohammed and D. Zeilberger. Sharp upper bounds for the orders of the recurrences output by the Zeilberger and q-Zeilberger algorithms. *J. Symbolic Comput.*, 39(2):201–207, 2005.
- [20] P. Paule. Greatest factorial factorization and symbolic summation. *J. Symbolic Comput.*, 20(3):235–268, 1995.
- [21] M. Petkovšek, H. S. Wilf, and D. Zeilberger. *A = B*. A. K. Peters Ltd., Wellesley, MA, 1996.
- [22] R. Pirastu and V. Strehl. Rational summation and Gosper-Petkovšek representation. *J. Symbolic Comput.*, 20(5-6):617–635, 1995.
- [23] C. Schneider. A new Sigma approach to multi-summation. *Adv. in Appl. Math.*, 34(4):740–767, 2005.
- [24] H. S. Wilf and D. Zeilberger. An algorithmic proof theory for hypergeometric (ordinary and “q”) multisum/integral identities. *Invent. Math.*, 108(3):575–633, 1992.
- [25] D. Zeilberger. A fast algorithm for proving terminating hypergeometric identities. *Discrete Math.*, 80(2):207–211, 1990.
- [26] D. Zeilberger. The method of creative telescoping. *J. Symbolic Comput.*, 11(3):195–204, 1991.

Appendix

In Theorem 3.3, we provide a criterion for the summability of rational functions of the form $\frac{a(x,y)}{d^j(x,y)}$, where $a \in \mathbb{Q}(x)[y]$ and $d \in \mathbb{Q}[x, y]$ is an irreducible polynomial. In this appendix, we present the proof of this criterion.

Firstly, we give the following lemma which proves the sufficiency of Theorem 3.3.

Lemma 4.5. *Let $f \in \mathbb{Q}(x, y)$ be of the form $f = \frac{a(x,y)}{d^j(x,y)}$, where $d \in \mathbb{Q}[x, y]$ is an irreducible polynomial, $a \in \mathbb{Q}(x)[y]$, and $\deg_y(a) < \deg_y(d)$. Suppose that there exist integers t, l with $t > 0$ and a polynomial $p(x, y) \in \mathbb{Q}(x)[y]$ such that*

$$\sigma_x^t d(x, y) = \sigma_y^l d(x, y),$$

and

$$a = \sigma_x^t \sigma_y^{-\ell} p(x, y) - p(x, y).$$

Then f is (σ_x, σ_y) -summable in $\mathbb{Q}(x, y)$.

Proof. Let $g = \sum_{k=0}^{t-1} \frac{\sigma_x^k(p)}{\sigma_x^k(d^j)}$, then

$$\frac{a}{d^j} - (\sigma_x g - g) = \frac{a}{d^j} - \frac{\sigma_x^t p}{\sigma_x^t d^j} + \frac{p}{d^j} = \frac{a+p}{d^j} - \frac{\sigma_x^t p}{\sigma_y^\ell d^j} = -\sigma_y^\ell \left(\frac{\sigma_x^t \sigma_y^{-\ell} p}{d^j} \right) + \frac{\sigma_x^t \sigma_y^{-\ell} p}{d^j}. \quad \blacksquare$$

The rest part of the appendix is devoted to proving the necessity of Theorem 3.3.

Analogue to the discrete residue given by Chen and Singer [10], we introduce the concept of polynomial residue. Let \mathbb{K} be a field and $f \in \mathbb{K}(x)$. By partial fraction decomposition, f can be written as

$$f = p(x) + \sum_{i=1}^m \sum_{j=1}^{n_i} \sum_{\ell=0}^{k_{i,j}} \frac{a_{i,j,\ell}(x)}{\sigma_x^\ell d_i^j(x)}, \quad (4.10)$$

where $p(x) \in \mathbb{K}[x]$, $m, n_i, k_{i,j} \in \mathbb{N}$, $\deg_x(a_{i,j,\ell}) < \deg_x(d_i)$, and $d_i(x)$ ($i = 1, \dots, m$) are irreducible polynomials that in distinct σ_x -orbits. The summation

$$\sum_{\ell=0}^{k_{i,j}} \sigma_x^{-\ell}(a_{i,j,\ell})$$

is called the *polynomial residue* of f at the σ_x -orbit of $d_i(x)$ of multiplicity j , denoted by $\text{res}_x(f(x), d_i(x), j)$.

It is easy to check that the summability of rational functions in $\mathbb{K}(x)$ can be given via polynomial residue. The proof is similar to the case of discrete residue [9, 10] and is omitted.

Proposition 4.6. *Let $f(x) \in \mathbb{K}(x)$ be of the form (4.10). Then $f(x)$ is σ_x -summable in $\mathbb{K}(x)$ if and only if the polynomial residue $\text{res}_x(f(x), d_i(x), j)$ is zero for any polynomial $d_i(x)$ and any multiplicity j .*

Now we are ready to prove the necessity of Theorem 3.3.

Suppose that $f = a/d^j$ is (σ_x, σ_y) -summable and assume that

$$f = \sigma_x g - g + \sigma_y h - h, \quad (4.11)$$

where $g, h \in \mathbb{Q}(x, y)$. As a univariate analogue to Lemma 3.1, we can decompose g into the form

$$g = \sigma_y g_1 - g_1 + g_2 + \frac{\lambda_1}{\sigma_x^{\mu_1} d^j} + \dots + \frac{\lambda_s}{\sigma_x^{\mu_s} d^j},$$

where $g_1, g_2 \in \mathbb{Q}(x, y)$ with g_2 containing no term of the form $\frac{\lambda}{\sigma_x^\mu d^j}$ in its partial fraction decomposition with respect to y , $\mu_\ell \in \mathbb{Z}$, $\lambda_\ell \in \mathbb{Q}(x)[y]$, and $\sigma_x^{\mu_\ell} d$ ($\ell = 1, \dots, s$) are irreducible polynomials in distinct σ_y -orbits.

Claim 1. Let

$$\Lambda := \{\sigma_x^{\mu_1} d, \dots, \sigma_x^{\mu_s} d, \sigma_x^{\mu_1+1} d, \dots, \sigma_x^{\mu_s+1} d\}.$$

Then

- (a) At least one element of Λ is in the same σ_y -orbit as d .
- (b) For each element $\eta \in \Lambda$, there is one element of $\Lambda \setminus \{\eta\} \cup \{d\}$ that is in the same σ_y -orbit as η .

Proof of Claim 1. (a) Suppose there is no element of Λ that is in the same σ_y -orbit as d . Since $f = a/d^j$, we have $\text{res}_y(f, d, j) = a \neq 0$. While by (4.11) and Proposition 4.6, we deduce that

$$\text{res}_y(f, d, j) = \text{res}_y(\sigma_x g - g, d, j) = 0,$$

which is a contradiction.

(b) The assertion follows from the same argument when considering the polynomial residues of η on both sides of (4.11). \blacksquare

Claim 1 implies that either $d \sim_y \sigma_x^{\mu'_1} d$ or $d \sim_y \sigma_x^{\mu'_1+1} d$ for some $\mu'_1 \in \{\mu_1, \dots, \mu_s\}$. We will only consider the first case. The second case can be treated similarly.

Claim 2. Assume $d \sim_y \sigma_x^{\mu'_1} d$. We have the following assertions.

(a) Suppose $k \geq 2$ be an integer such that $\sigma_x^l d \sim_y d$ for $1 \leq l \leq k-1$. Then there exist $\mu'_1, \dots, \mu'_k \in \{\mu_1, \dots, \mu_s\}$ such that

$$\sigma_x^{\mu'_1+1} d \sim_y \sigma_x^{\mu'_2} d, \quad \sigma_x^{\mu'_2+1} d \sim_y \sigma_x^{\mu'_3} d, \quad \dots, \quad \sigma_x^{\mu'_{k-1}+1} d \sim_y \sigma_x^{\mu'_k} d,$$

and

$$\sigma_x^{k-1} d \sim_y \sigma_x^{\mu'_k} d.$$

(b) There exists a positive integer $t \leq s$ such that $\sigma_x^t d \sim_y d$.

Proof of Claim 2. (a) From Claim 1(b), we derive that $\sigma_x^{\mu'_1+1} d$ is σ_y -equivalent to an element of $\Lambda \setminus \{\sigma_x^{\mu'_1+1} d\} \cup \{d\}$. If $\sigma_x^{\mu'_1+1} d \sim_y d$, then $\sigma_x^{\mu'_1+1} d \sim_y \sigma_x^{\mu'_1} d$ and thus $\sigma_x d \sim_y d$, which contradicts to the hypothesis on k . If $\sigma_x^{\mu'_1+1} d \sim_y \sigma_x^{\mu'_l+1} d$, then $\sigma_x^{\mu'_1} d \sim_y \sigma_x^{\mu'_l} d$ for some l , which contradicts to the assumption that $\sigma_x^{\mu'_i}$ are in distinct σ_y -orbits. Therefore we are left with the only possibility that $\sigma_x^{\mu'_1+1} d \sim_y \sigma_x^{\mu'_2} d$ for some $\mu'_2 \in \{\mu_1, \dots, \mu_s\} \setminus \{\mu'_1\}$. Continue this process, we will find μ'_3, \dots, μ'_k such that

$$\sigma_x^{\mu'_2+1} d \sim_y \sigma_x^{\mu'_3} d, \quad \dots, \quad \sigma_x^{\mu'_{k-1}+1} d \sim_y \sigma_x^{\mu'_k} d.$$

Finally, we have

$$\sigma_x^{\mu'_k} d \sim_y \sigma_x^{\mu'_1+k-1} d \sim_y \sigma_x^{k-1} d.$$

(b) If such t does not exist, then one could find $\{\mu'_1, \dots, \mu'_{s+1}\}$ satisfying the constraints in (a). Thus, it holds that $\mu'_r = \mu'_t$ for some $r > t$. Hence $\sigma_x^{\mu'_1+r} d \sim_y \sigma_x^{\mu'_1+t} d$, which leads to $\sigma_x^{r-t} d \sim_y d$, a contradiction. \blacksquare

Suppose t is the smallest integer such that $\sigma_x^t d \sim_y d$. Then taking $k = t$ in Claim 2(a), we derive that there exist $\mu'_1, \dots, \mu'_t \in \{\mu_1, \dots, \mu_s\}$ such that

$$\sigma_x^{\mu'_1+1} d \sim_y \sigma_x^{\mu'_2} d, \quad \sigma_x^{\mu'_2+1} d \sim_y \sigma_x^{\mu'_3} d, \quad \dots, \quad \sigma_x^{\mu'_{t-1}+1} d \sim_y \sigma_x^{\mu'_t} d,$$

and

$$\sigma_x^{\mu'_t+1} d \sim_y \sigma_x^t d \sim_y d.$$

Recall that $\sigma_x^{\mu'_1} d \sim_y d$. By the definition of \sim_y , there exist integers s_0, s_1, \dots, s_t such that

$$\sigma_x^{\mu'_k+1} d = \sigma_x^{\mu'_{k+1}} \sigma_y^{s_{k+1}} d, \quad 1 \leq k \leq t-1, \quad \sigma_x^{\mu'_t+1} d = \sigma_y^{s_1} d, \quad \text{and} \quad \sigma_x^{\mu'_1} d = \sigma_y^{s_0} d.$$

Hence,

$$\sigma_y^{s_0} d = \sigma_x^{\mu'_1} d = \sigma_x^{\mu'_2-1} \sigma_y^{s_2} d = \sigma_x^{\mu'_3-2} \sigma_y^{s_2+s_3} d = \dots = \sigma_x^{\mu'_t-t+1} \sigma_y^{s_2+s_3+\dots+s_t} d = \sigma_y^{s_1+\dots+s_t} \sigma_x^{-t} d.$$

Setting $\ell = s_1 + \dots + s_t - s_0$, we then have $\sigma_x^t d = \sigma_y^\ell d$.

Now we compare the polynomial residues on both sides of (4.11). We list the residues

σ_y -orbit	Comparison of two sides of (4.11)
$d, \sigma_x^{\mu'_t+1} d$	$a = \sigma_x \sigma_y^{-s_0} \lambda'_t - \sigma_y^{-s_1} \lambda'_1$
$\sigma_x^{\mu'_{t-1}+1} d, \sigma_x^{\mu'_t} d$	$0 = \sigma_x \sigma_y^{-s_t} \lambda'_{t-1} - \lambda'_t$
$\sigma_x^{\mu'_{t-2}+1} d, \sigma_x^{\mu'_{t-1}} d$	$0 = \sigma_x \sigma_y^{-s_{t-1}} \lambda'_{t-2} - \lambda'_{t-1}$
\vdots	\vdots
$\sigma_x^{\mu'_2+1} d, \sigma_x^{\mu'_3} d$	$0 = \sigma_x \sigma_y^{-s_3} \lambda'_2 - \lambda'_3$
$\sigma_x^{\mu'_1+1} d, \sigma_x^{\mu'_2} d$	$0 = \sigma_x \sigma_y^{-s_2} \lambda'_1 - \lambda'_2$

Table 1: Orbits and their corresponding polynomial residues.

in Table 1, where the first column consists of the σ_y -orbits of elements in Λ and the second column consists of the equations obtained by equating the corresponding polynomial residues on both sides of (4.11). By investigating the equations in Table 1 from bottom to top, we find that

$$a = \sigma_x^t \sigma_y^{-\ell} p - p,$$

where $p = \sigma_y^{-s_1} \lambda'_1(x, y)$. Since $\deg_y \lambda'_1 < \deg_y d$, we have $\deg_y p < \deg_y d$. This completes the proof of Theorem 3.3. \blacksquare